Technical Report 1126

Analog Scales as Temperament Measures in the Baseline Officer Longitudinal Data Set (BOLDS)

Lynn M. Milan
U. S. Army Research Institute

MARCH 2002

**United States Army Research Institute
for the Behavioral and Social Sciences**

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) March 2002	2. REPORT TYPE Interim	3. DATES COVERED (from. . . to) July 2000 – July 2001

4. TITLE AND SUBTITLE Analog Scales as Temperament Measures in the Baseline Officer Longitudinal Data Set (BOLDS)	5a. CONTRACT OR GRANT NUMBER
	5b. PROGRAM ELEMENT NUMBER 20262785

6. AUTHOR(S) Lynn M. Milan (U.S. Army Research Institute)	5c. PROJECT NUMBER A790
	5d. TASK NUMBER WP103
	5e. WORK UNIT NUMBER H01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: TAPC-ARI-RK 5001 Eisenhower Avenue Alexandria, VA 22333-5600	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Technical Report 1126

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

The Baseline Officer Longitudinal Data Set (BOLDS) was developed jointly by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and the U.S. Military Academy (USMA) to enable researchers to study the development of leader performance over time. Currently, BOLDS consists of data accumulated on USMA cadets from the Class of 1998. Because no personality inventories were administered to cadets, analog scales were developed to represent the scales of the *Assessment of Background and Life Experiences* (ABLE) and the *NEO Personality Inventory* (NEO-PI). This report examines these empirically derived analog scales to ensure they measure within BOLDS what they were intended to measure. Two replication exercises were performed: (1) rerunning analyses Evans (1997) ran when creating the analog scales and (2) comparing results produced using the analog scales with those that emerged from *actual* ABLE scales. Results indicate the analogs perform in a manner consistent with previous analyses, suggesting they suffice as veridical temperament measures and may be used as such in future BOLDS analyses.

15. SUBJECT TERMS
Analog scales, temperament measures, personality assessment, USMA cadets, leader attributes, longitudinal database

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	41	Lynn M. Milan 703-617-0317

Analog Scales as Temperament Measures in the Baseline Officer Longitudinal Data Set (BOLDS)

Lynn M. Milan
U. S. Army Research Institute

Leader Development Research Unit
Stanley M. Halpin, Chief

March 2002

FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducts research on personnel performance and training in support of Army goals. A primary concern of the U.S. Army is developing effective future leaders. To advance research on leader development and enhance our understanding of the leader development process, ARI and the United States Military Academy (USMA) initiated a cooperative effort to examine the development of leaders from a longitudinal perspective.

The initial phase of this effort collected information on a variety of personal attributes and performance measures from USMA cadets in the Class of 1998, creating the Baseline Officer Longitudinal Data Set (BOLDS). This report examines temperament scales that were empirically derived from archival survey data and included in BOLDS to estimate various personality constructs in cadets. Results were reported to the Leader Development Research Center at USMA.

Because future phases of the BOLDS project will continue to track these officers through their Army careers, an understanding of the measures available in the database will be important as researchers investigate the development of these individuals, in attempts to improve leader development efforts for the Objective Force.

MICHAEL G. RUMSEY
Acting Technical Director

ACKNOWLEDGMENT

ANALOG SCALES AS TEMPERAMENT MEASURES IN THE BASELINE OFFICER LONGITUDINAL DATA SET (BOLDS)

EXECUTIVE SUMMARY

Research Requirement:

Due to time constraints, personality measures were not administered to United States Military Academy (USMA) cadets in the Class of 1998 as part of the Baseline Officer Longitudinal Data Set (BOLDS). To compensate for this data deficit, Evans (1997) developed analog scales to empirically mimic scales of the "Assessment of Background and Life Experiences" (ABLE; White, Nord, Mael, & Young, 1993) and the "NEO Personality Inventory" (NEO-PI; Costa & McCrae, 1985). Cadets' scores on these analog scales are included in BOLDS. This paper sought to examine whether the analog scales reliably measure what they purport to measure and, hence, whether they should be used in analyses of BOLDS.

Procedure:

Two replication exercises were undertaken to explore whether the analog scales in BOLDS performed comparably to previous analyses. First, replicating the analyses conducted by Evans (1997) during analog scale development, analog scale scores of BOLDS cadets were correlated with their leadership grades, and the scores of USMA graduates were compared with those of non-graduates. Second, replicating analyses conducted by Mael and White (1994), the analog scale scores from BOLDS were correlated with leadership criterion measures and with leadership predictors used by USMA. In addition, hierarchical regressions were performed to explore whether the analog ABLE Total score accounts for significant variance over and above the score USMA has used in the past to predict cadet success.

Findings:

        The vast majority (five-sixths) of the correlations between analog ABLE/NEO-PI scale scores and cadets' leadership grades showed no significant difference between the correlations resulting from BOLDS and those from Evans' initial data. Similarly, three-fourths of the t-tests comparing the scores of USMA graduates with those of non-graduates produced the same results across the two datasets.

        Correlations of the BOLDS analog ABLE scores with leadership criterion measures and leadership predictors resembled the correlations Mael and White (1994) reported for the actual ABLE scales. Likewise, results from the hierarchical regressions including the analog ABLE Total score reproduced Mael and White's finding that ABLE Total provides incremental validity over and above the score USMA typically uses to predict cadet success.

Utilization of Findings:

        Because the analog ABLE/NEO-PI scales performed similarly using BOLDS data as they had in prior analyses, they appear to be stable measures. Therefore, in analyzing BOLDS data, these analog scales may be useful as measures of cadet temperament.

ANALOG SCALES AS TEMPERAMENT MEASURES IN THE BASELINE OFFICER LONGITUDINAL DATA SET (BOLDS)

CONTENTS

CONTENTS (continued)

**ANALOG SCALES AS TEMPERAMENT MEASURES IN THE
BASELINE OFFICER LONGITUDINAL DATA SET (BOLDS)**

INTRODUCTION

Beginning in 1993, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) collaborated with the United States Military Academy (USMA) to create a database of information about cadets in USMA's Class of 1998. The intent of the project was to develop a longitudinal database that would enable researchers to examine changes in individual leaders' performance and effectiveness over time. The resulting *Baseline Officer Longitudinal Data Set* (BOLDS) contains measures of cognitive aptitude, problem-solving skills, tacit knowledge of military leadership, motivation, leadership style, leadership performance, physical fitness, cognitive-emotional development, and developmental experiences (for details about these measures, see Milan, Bourne, Zazanis, & Bartone, in press).

BOLDS also includes measures of temperament, but – aside from *hardiness* – these variables are based on empirically derived estimates (known as *analog scales*) of personality constructs rather than direct measures. BOLDS researchers recognized that temperament measures could be useful predictors in their longitudinal research, given that certain personality dimensions have been found to be associated with leadership effectiveness (see Hogan, Curphy, & Hogan, 1994; Zaccaro, 1996). Nonetheless, due to the limited amount of cadet time available, they decided to focus on administering measures uniquely associated with leadership. Therefore, no personality measures were administered directly to cadets in the Class of 1998 while they were at West Point.

As part of separate research programs, however, USMA cadets in prior classes had completed two personality measures. The Class of 1994 completed a short form of the Assessment of Background and Life Experiences (ABLE; White, Nord, Mael, & Young, 1993), an instrument developed to predict successful performance of U.S. Army jobs. This short form consisted of five content scales (Dominance, Energy Level, Work Orientation, Emotional Stability, and Traditional Values) and one response validity scale (Social Desirability; see Table 1 for scale descriptions). Likewise, the Class of 1996 completed the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985), which measures the five major dimensions of normal adult personality (Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness; see Table 2).

To enhance the BOLDS database on the Class of 1998, Evans (1997) constructed analog scales to replicate these ABLE and NEO-PI scales. To create the ABLE analog scales, archival data was needed that was available for both the Class of 1998 and the Class of 1994. Similarly, to create the NEO-PI analog scales, archival data was needed that was available for both the Class of 1998 and the Class of 1996.

Such archival data existed as a result of two self-measures administered annually to the incoming class of cadets. The first of these, the Student Information Form (SIF), is a national survey completed by entering freshman at hundreds of universities. It is part of the Cooperative

Institutional Research Program (CIRP) sponsored by the American Council on Education and the University of California at Los Angeles (Astin, Korn, & Berz, 1990). Items on the SIF ask about high school background, college expectations, family characteristics, personal attributes, attitudes, and goals. The USMA Class of 1994 completed the 1990 Student Information Form (Astin et al., 1990), the Class of 1996 completed the 1992 SIF, and the Class of 1998 completed the 1994 SIF. The second measure from which archival data were obtained is the Class Characteristics Inventory (CCI), developed by the Institutional Research and Analysis Branch (IRAB) of USMA's Office of Policy, Planning, and Analysis. The CCI describes members of each entering class in terms of high school experiences (academic, extracurricular, athletic), admission to USMA, and personal characteristics. Though both surveys are revised annually to meet changing research needs, their items remain fairly consistent from year to year.

For each ABLE and NEO-PI scale, Evans found the 20 archival survey items having the highest zero-order correlations with the scale. A series of multiple regression analyses were then run to predict each scale score. For purposes of double cross-validation, the predictors remaining in the final stepwise regression model were used in a simultaneous regression analysis performed on each random half of the sample. Mean parameter estimates calculated from the results of the two random halves were used to compute scale analogs (see Appendixes A and B for the prediction equations). Lastly, the relationships of the original scales and their analogs to leadership grades and cadet attrition were compared. (The preceding brief methodological description is intended to provide a sense of Evans' empirical approach to developing analog scales; for details, see Evans, 1997.)

Evans (1997) concluded that the analog scales were "reasonably close approximations" of the original scales because they accounted for a substantial proportion of the variance in the original scales and they exhibited similar relationships with criterion measures. Consequently, he recommended the use of the analog scales in future BOLDS analyses.

Table 1.
Brief Description of the ABLE Scales

ABLE Scale	Related Construct
Dominance	Surgency: The tendency to seek and enjoy positions of leadership and influence over others
Energy Level	Surgency: The amount of energy and enthusiasm a person has
Work Orientation	Achievement: The tendency to strive for competency in one's work
Emotional Stability	Adjustment: The amount of emotional stability and tolerance for stress a person possesses
Traditional Values	Dependability: A person's acceptance of societal values
ABLE Total[1]	Broad concept of adaptability[2]
Social Desirability	Intentional distortion of self-descriptions in a favorable direction

Source: Hough et al., 1990.
[1] ABLE Total was a composite of items from the other five temperament scales (i.e., not including the Social Desirability scale).
[2] According to White, Nord, Mael, & Young, 1993.

Table 2.
Brief Description of the NEO-PI Scales

NEO-PI Scale	Related Facets
Neuroticism (versus emotional stability)	Anxiety, angry hostility, depression, self-consciousness, impulsiveness, vulnerability
Extraversion (versus introversion)	Gregariousness, assertiveness, activity, excitement-seeking, positive emotions, warmth
Openness to experience (versus closedness)	Ideas, fantasy, aesthetics, actions, feelings, values
Agreeableness (versus antagonism)	Trust, straightforwardness, altruism, compliance, modesty, tender-mindedness
Conscientiousness (versus lack of direction)	Competence, order, dutifulness, achievement striving, self-discipline, deliberation

Source: John & Srivastava, 1999.

PART I: REPLICATING EVANS' (1997) ANALYSES REGARDING THE ANALOG ABLE AND ANALOG NEO-PI SCALES

The current effort sought to replicate Evans' analyses associating the analog scale scores with criterion measures (i.e., leadership grades and cadet attrition), this time using BOLDS data. In other words, the analog scale scores of the BOLDS cadets were correlated with their leadership grades, and the scores of graduates were compared with those of non-graduates – just as Evans had done. All results were compared to those reported by Evans (1997). The immediate objective was to examine whether the analog scales would perform similarly within the BOLDS data as they had within the data from which they were developed. The ultimate goal was to recommend whether the analog ABLE and analog NEO-PI scales should be used in future BOLDS analyses.

Method

Participants

USMA Class of 1994

Of the 1,325 incoming cadets in the Class of 1994, 727 completed the ABLE and responded to each of the 52 predictor items initially selected from the SIF and the CCI. To determine the representativeness of these 727 cadets, Evans (1997) compared their ABLE scale means and scale intercorrelations to those of the entire class. Because these statistics were "extremely close," Evans concluded that the 727-cadet sample was representative of the Class of 1994, in terms of its ABLE scale characteristics.

USMA Class of 1996

Of the 1,049 cadets in the Class of 1996 who completed the NEO-PI during the summer of 1992, 635 also responded to 76 predictor items drawn from the SIF and the CCI administered at the same time. Evans (1997) compared the NEO-PI scale mean scores and scale intercorrelations from the 635 cadets to the scale statistics from the entire class and concluded the 635-cadet sample was representative in terms of its NEO-PI scale characteristics.

USMA Class of 1998

As shown in Table 3, 1143 cadets attended Cadet Basic Training (CBT) in the summer of 1994 prior to their freshman year, and 1006[1] remained to complete their first academic semester. By the middle of their senior year, 859 cadets remained in the Class of 1998.[2] In the analyses

---

[1] Of this group, 87% were male. Similarly, the 635-cadet sample from the Class of 1996 was composed of 88% males. The gender composition of the sample from the Class of 1994 was not available.

[2] These figures are based on the number of cadets in the BOLDS database who had a final Military Development (MD) grade for the term. It is acknowledged that a few MD grades each term could be missing from BOLDS, thus under-representing the true number of cadets remaining in the class. In actuality, of the 1143 cadets who were admitted in 1994, 883 (77%) graduated in the Class of 1998.

that follow, "graduates" were considered to be those cadets who remained after their seventh semester at USMA (i.e., $N = 859$), and "non-graduates" were considered to be those individuals who left USMA at any time prior to completing their seventh term ($N = 284$).

Table 3.
Samples and Sample Sizes From the BOLDS Data

Class of 1998 Sample	Size
Cadets attending CBT in the Summer of 1994	1143
Cadets who completed their first academic semester at USMA (Fall 1994)	1006
Graduating cadets in the Class of 1998	859[1]
Non-graduates originally admitted in the Class of 1998	284[1]
• Left prior to completing first semester	137
• Left after first semester	147

[1] Based on 7th semester data.

Measures

Analog Scales

The regression equations Evans developed to calculate analog scores for the seven ABLE scales included a total of 35 different predictor items, 30 from the SIF and 5 from the CCI. Of these predictors, 1 was used to estimate five scales, 5 were used to estimate four scales, 9 were used to estimate three scales, 4 were used to estimate two scales, and 16 were used to estimate only one scale (see Appendix A for the prediction equations).

Similarly, the regression equations representing the five NEO-PI scales included a total of 47 different predictors, 45 from the SIF and 2 from the CCI. Of these predictors, 1 was used to estimate three scales, 8 were used to estimate two scales, and 38 were used to estimate only one scale (see Appendix B for the actual equations).

In keeping with Evans' formulation, BOLDS researchers computed analog ABLE and NEO-PI scale scores for each cadet who was admitted to USMA in the summer of 1994 and completed the SIF and CCI. Because each analog scale is composed of 9 to 13 survey items, scores for all the scales could not be computed for each cadet, due to missing data. For each scale, scores are missing for 17-36% of cadets.

External Criteria

At the end of each academic semester, cadets received a leadership (or "military development") grade (MD), which is the main evaluative measure of a cadet's military

performance for that term. The leadership grade is based on a conventional five-point scale: $A(4) - B(3) - C(2) - D(1) - F(0)$. It is determined by calculating a weighted average of the grades assigned by various raters. For cadets in most duty positions,[3] 50% of their grade is determined by their Tactical Officer,[4] 30% comes from their immediate superior in the cadet chain of command, 10% is determined by their second-level cadet superior, and the final 10% comes from their third-level cadet superior. An exception occurs when a Tactical Officer gives a grade of D or F, in which case this grade forms 100% of the cadet's final grade. In keeping with a forced distribution system, no more than 20% of the cadets graded *by an individual* can receive an A, no more than 40% can receive a B, and no more than 40% can receive a C. The awarding of Ds or Fs is not limited, though these grades rarely occur.

In addition to the academic semesters, cadets also participate in summer details, activities intended to provide cadets with hands-on military training in a field environment. As with the semesters, at the end of each summer detail, cadets received a leadership grade (MD). Because the leadership grades from the academic semesters and the summer details may represent distinct leadership performance dimensions (Mael & Hirsch, 1993; Mael & White, 1994), Evans examined the two sets of grades separately.

Results

ABLE

Table 4 shows the correlations between the ABLE scale scores and the ABLE analog scale scores with the mean leadership grades that USMA graduates earned during their academic semesters and summer details. As a general observation, the relationships between mean leadership grades and the analog scales in the BOLDS data were slightly stronger than the relationships using the original ABLE scales and substantially stronger than the relationships using the ABLE analog scales with data from the Class of 1994. Still, none of the correlations between BOLDS analog scales and MD grades was particularly strong, ranging from .07 to .17.

Across all three measurement techniques, three scales were consistently correlated with the academic semester leadership grade: Work Orientation, Social Desirability, and Traditional Values, with Work Orientation showing the largest magnitude (.17 – .26). In fact, Work Orientation and Social Desirability were also consistently correlated with the summer leadership grade. (The correlation between mean academic semester MD grade and mean summer MD grade for graduating cadets was .56.) Regarding inconsistencies, the analog scale scores computed on 1994 data often showed no correlation with leadership grades, unlike the original ABLE scale scores and the analog scale scores calculated from BOLDS data (e.g., Dominance, Energy Level, Total).

---

[3] Duty positions are cadets' job assignments within the military hierarchy of the U.S. Corps of Cadets. Duty positions differ in their functional task requirements and their level of supervisory responsibility. Assignments are made based on cadets' graduation year and their past performance.

[4] A Tactical Officer is a military officer assigned to a unit of cadets and charged with monitoring and developing the cadets.

To determine whether the correlations between the BOLDS analog scale scores and mean leadership grades differed from the correlations using the 1994 analog scale scores, Fisher's z' transformations and comparisons between independent rs were computed. In four of the fourteen cases, the correlations differed. Three of the differences occurred in relation to the mean grades for summer details and the scales of Emotional Stability, Dominance, and the ABLE Total. In each of these three instances, the BOLDS data showed a significant positive correlation, whereas the 1994 data indicated essentially no correlation (i.e., $r = .03 - .05$). The fourth case involved the correlation between the Emotional Stability scale score and the mean MD grade from the academic semesters. From the BOLDS data, a very weak positive correlation emerged ($r = .07$), whereas a very weak negative correlation ($r = -.06$) appeared in the 1994 data.

Table 5 compares USMA graduates and non-graduates on their mean scores from the ABLE scales and ABLE analog scales. In all cases, graduates' mean scores were higher than those of non-graduates. Though the magnitude of the differences was relatively small, in all but three instances the differences were statistically significant. In two cases involving the original ABLE scales (Work Orientation and Social Desirability) and one case involving the BOLDS data (Dominance Scale), the differences in mean scores were not statistically significant. Another interesting observation, though its interpretation is unclear, is BOLDS means were lower than the 1994 means in all instances but one (i.e., Dominance among non-graduates).

From the perspective of selecting and retaining quality cadets, one question that arises is whether the individuals who left right after CBT or prior to completing their first academic semester differ in some way from cadets who left later in their USMA career. For instance, individuals who attrit immediately may do so because of motivational factors, whereas those who attrit later may do so because of weaknesses in knowledge, skills, or abilities (KSAs). If USMA could identify potential immediate attriters, admissions officers could attempt to prepare them better for the USMA experience, thereby increasing the probability that cadets who choose to attend are motivationally committed to graduating. Regarding the analog ABLE scale scores, t-tests revealed no significant differences between the mean scores of these two groups of non-graduates (see Table 6).

Table 4.
Correlations of ABLE/Analog ABLE Scale Scores With Mean Leadership Grades
for Graduating Cadets

ABLE Scale	N	Correlation with Mean Leadership (MD) Grades	
		Academic Semesters[1]	Summer Details[2]
'94 ABLE Dominance	553	.09*	.09*
'94 Analog Dominance	553	.06	**.03**
'98 BOLDS Analog Dominance	675	.14***	**.16****
z		-1.41	-2.28*
'94 ABLE Energy Level	553	.08*	.13**
'94 Analog Energy Level	553	.04	.06
'98 BOLDS Analog Energy Level	629	.15***	.15***
z		-1.90	-1.56
'94 ABLE Work Orientation	553	.26****	.13**
'94 Analog Work Orientation	553	.22****	.08*
'98 BOLDS Analog Work Orientation	709	.17****	.11**
z		0.91	-0.53
'94 ABLE Emotional Stability	553	-.02	.06
'94 Analog Emotional Stability	553	**-.06**	**.03**
'98 BOLDS Analog Emotional Stability	595	**.07**	**.15****
z		-2.20*	-2.04*
'94 ABLE Traditional Values	553	.15***	.10*
'94 Analog Traditional Values	553	.14**	.06
'98 BOLDS Analog Traditional Values	550	.17****	.15***
z		-0.51	-1.51
'94 ABLE Total	553	.15***	.14***
'94 Analog Total	553	.05	**.05**
'98 BOLDS Analog Total	595	.13**	**.17****
z		-1.37	-2.06*
'94 ABLE Social Desirability	553	.13**	.14***
'94 Analog Social Desirability	553	.17****	.11*
'98 BOLDS Analog Social Desirability	610	.12**	.11**
z		0.87	0.00

Note. Boldface type indicates a statistically significant difference between the correlations resulting from the two applications of the *analog* scales.
[1]For the '94 data, mean was based on the leadership grades earned over eight semesters. For the BOLDS data, mean includes grades earned during academic semesters 1-7 (8th semester grade is not included).
[2]For the '94 data, mean was based on the leadership grades earned during four summer details. For the BOLDS data, mean includes grades from two details during the summer of 1995, two details during the summer of 1996, and one detail during the summer of 1997.
*$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$. (Four significance levels are distinguished to replicate Evans, 1997.)

Table 5.
A Comparison of the Mean ABLE/Analog ABLE Scale Scores of
Graduates and Non-Graduates

ABLE Scale	Graduates		Non-Graduates			
	M	*SD*	*M*	*SD*	*t*	*df*
'94 ABLE Dominance	2.56	.31	2.49	.34	2.55*	725.0
'94 Analog Dominance	2.56	.23	2.51	.26	2.09*	264.6
'98 BOLDS Analog Dominance	2.53	.23	2.52	.27	0.43	301.1
'94 ABLE Energy Level	2.38	.27	2.29	.32	3.41***	257.3
'94 Analog Energy Level	2.37	.17	2.30	.22	4.07***	247.2
'98 BOLDS Analog Energy Level	2.34	.18	2.28	.22	3.86***	288.1
'94 ABLE Work Orientation	2.39	.34	2.34	.40	1.53	259.1
'94 Analog Work Orientation	2.40	.24	2.32	.28	3.18**	255.4
'98 BOLDS Analog Work Orientation	2.34	.24	2.29	.27	2.86**	923.0
'94 ABLE Emotional Stability	2.39	.28	2.29	.31	4.05***	725.0
'94 Analog Emotional Stability	2.38	.19	2.31	.22	4.22***	258.6
'98 BOLDS Analog Emotional Stability	2.35	.19	2.28	.21	4.13***	768.0
'94 ABLE Traditional Values	2.57	.28	2.49	.33	3.13**	255.9
'94 Analog Traditional Values	2.57	.18	2.52	.18	3.34***	725.0
'98 BOLDS Analog Traditional Values	2.54	.18	2.48	.22	3.21**	228.8
'94 ABLE Total	2.44	.20	2.36	.24	4.11***	255.9
'94 Analog Total	2.44	.15	2.37	.19	4.38***	247.6
'98 BOLDS Analog Total	2.41	.16	2.35	.18	3.93***	768.0
'94 ABLE Social Desirability	1.42	.24	1.39	.24	1.29	725.0
'94 Analog Social Desirability	1.42	.12	1.38	.13	3.79***	725.0
'98 BOLDS Analog Social Desirability	1.40	.12	1.36	.13	4.21***	799.0

Note. 1994: Graduates *N* = 553 and Non-graduates *N* = 174. 1998: Graduates *N* = 590-709 and Non-graduates *N* = 160-216.
*p < .05. **p < .01. ***p < .001. (Only three significance levels could be distinguished based on the p-values cited by Evans, 1997.)

Table 6.
A Comparison of the Mean Analog ABLE Scale Scores of Non-Graduates

ABLE Scale	Time 1 Attriters		Time 2 Attriters			
	M	*SD*	*M*	*SD*	*t*	*df*
BOLDS Analog Dominance	2.52	.27	2.52	.26	-0.01	203
BOLDS Analog Energy Level	2.25	.22	2.30	.21	1.54	196
BOLDS Analog Work Orientation	2.29	.29	2.28	.25	-0.25	214
BOLDS Analog Emotional Stability	2.27	.19	2.30	.21	1.03	173
BOLDS Analog Traditional Values	2.46	.22	2.50	.22	1.14	158
BOLDS Analog Total	2.33	.17	2.37	.18	1.22	173
BOLDS Analog Social Desirability	1.35	.14	1.36	.12	0.50	189

Note. "Time 1 Attriters" refers to cadets who left USMA prior to completing their first academic semester ($N = 70$-107), and "Time 2 Attriters" refers to those who left anytime after their first semester ($N = 90$-109). No *t*-values were statistically significant.

NEO-PI

The same comparisons made using the analog ABLE scales were examined using the analog NEO-PI scales. Correlations between cadets' NEO-PI scale scores and mean leadership grades are shown in Table 7. All three Conscientiousness Scale scores (i.e., original NEO-PI, 1996 analog, and BOLDS analog) were positively correlated with cadets' mean leadership grades from both the academic semesters and the summer details. (In fact, of all the NEO-PI scales, Conscientiousness showed the highest correlations with MD grades, ranging from .15 to .29.) All three Agreeableness scores were positively correlated with the mean academic semester grade, and all three Extraversion scores were positively correlated with the mean summer detail grade.

Three minor inconsistencies also appeared. The original NEO-PI Openness Scale was negatively correlated with the academic semester mean leadership grade, and the original NEO-PI Agreeableness Scale was positively correlated with the summer detail mean grade. In both cases, the analog scale scores (from both 1996 and BOLDS) exhibited no significant correlation. In contrast, both analog Neuroticism Scale scores were negatively correlated with the summer detail grade, whereas the original Neuroticism Scale score showed no significant correlation.

To examine whether the correlations resulting from the two applications of the analog scales (1996 and BOLDS) differed, Fisher's *z'* transformations and comparisons between independent *r*s were calculated. As shown in Table 7, none of the pairs contained correlations that significantly differed from one another; in other words, the analog scales produced consistent results across two data sets.

Regarding the *t*-tests conducted to determine whether graduates and non-graduates differed in their NEO-PI scale scores, there seemed to be little consistency, other than the fact that in all three cases, graduates and non-graduates scored similarly on the Openness scale (see Table 8). For two scales (Neuroticism and Conscientiousness), the original NEO-PI scale scores showed no difference between graduates and non-graduates, whereas the analog scales indicated a significant difference both in 1996 and among BOLDS cadets. With respect to two other scales, the results from the BOLDS data differed from the 1996 findings. Specifically, the BOLDS data showed no difference between graduates and non-graduates on the Extraversion Scale, while the original Extraversion Scale scores and the 1996 analog scores demonstrated a difference. Conversely, the BOLDS data suggested a difference between graduates and non-graduates on Agreeableness, whereas the 1996 findings (both original scale and analog) showed no difference.

Among non-graduates, those who left immediately after CBT or before completing their first academic semester at USMA were compared to cadets who left later, with regard to their analog NEO-PI scales scores (see Table 9). According to *t*-test results, no statistically significant differences in mean scale scores exist between these two groups of non-graduates.

11

Table 7.
Correlations of NEO-PI/Analog NEO-PI Scale Scores With Mean Leadership Grades for Graduating Cadets

NEO-PI Scale	N	Correlation with Mean Leadership (MD) Grades	
		Academic Semesters[1]	Summer Details[2]
'96 NEO-PI Neuroticism	496	-.04	-.09
'96 Analog Neuroticism	496	-.06	-.12**
'98 BOLDS Analog Neuroticism	674	-.09*	-.12**
z		0.51	0.00
'96 NEO-PI Extraversion	496	.04	.16***
'96 Analog Extraversion	496	.02	.16***
'98 BOLDS Analog Extraversion	652	.11**	.17****
z		-1.51	-0.18
'96 NEO-PI Openness	496	-.09*	-.06
'96 Analog Openness	496	-.02	-.05
'98 BOLDS Analog Openness	676	-.02	.02
z		0.00	-1.18
'96 NEO-PI Agreeableness	496	.14**	.10*
'96 Analog Agreeableness	496	.13**	.06
'98 BOLDS Analog Agreeableness	692	.10**	.02
z		0.53	0.68
'96 NEO-PI Conscientiousness	496	.29****	.22****
'96 Analog Conscientiousness	496	.19****	.22****
'98 BOLDS Analog Conscientious.	591	.15***	.15***
z		0.67	1.20

[1]For the '94 data, mean was based on the leadership grades earned over eight semesters. For the BOLDS data, mean includes grades earned during academic semesters 1-7 (8th semester grade is not included).

[2]For the '94 data, mean was based on the leadership grades earned during four summer details. For the BOLDS data, mean includes grades from two details during the summer of 1995, two details during the summer of 1996, and one detail during the summer of 1997.

*$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$. (Four significance levels are distinguished to replicate Evans, 1997.)

Table 8.
A Comparison of the Mean NEO-PI/Analog NEO-PI Scale Scores of
Graduates and Non-Graduates

NEO-PI Scale	Graduates		Non-Graduates			
	M	*SD*	*M*	*SD*	*t*	*df*
'96 NEO-PI Neuroticism	90.76	21.73	93.04	25.01	0.97	200.1
'96 Analog Neuroticism	89.90	14.37	95.54	17.06	3.56***	196.2
'98 BOLDS Analog Neuroticism	90.09	13.83	96.68	15.71	-5.79***	877.0
'96 NEO-PI Extraversion	120.21	17.71	115.93	19.38	-2.47*	633.0
'96 Analog Extraversion	119.85	11.31	116.81	11.76	-2.71**	633.0
'98 BOLDS Analog Extraversion	119.62	10.56	119.60	12.76	.01	283.8
'96 NEO-PI Openness	111.93	17.48	113.54	17.63	0.96	633.0
'96 Analog Openness	112.42	11.52	112.12	11.22	-.28	633.0
'98 BOLDS Analog Openness	110.49	12.39	111.42	13.09	-.93	878.0
'96 NEO-PI Agreeableness	108.38	17.29	106.22	18.33	-1.28	633.0
'96 Analog Agreeableness	108.00	9.08	107.35	9.23	-0.74	633.0
'98 BOLDS Analog Agreeableness	107.53	8.65	105.75	8.54	2.61**	899.0
'96 NEO-PI Conscientiousness	117.32	19.42	113.73	21.34	-1.89	633.0
'96 Analog Conscientiousness	117.50	12.72	112.82	13.86	-3.75***	633.0
'98 BOLDS Analog Conscientiousness	116.77	13.06	114.09	13.98	2.36*	766.0

Note. 1996: Graduates $N = 496$ and Non-graduates $N = 139$. 1998: Graduates $N = 591-692$ and Non-graduates $N = 177-209$.
*$p < .05$. **$p < .01$. ***$p < .001$. (Only three significance levels could be distinguished based on the *p*-values cited by Evans, 1997.)

Table 9.
A Comparison of the Mean Analog NEO-PI Scale Scores of Non-Graduates

NEO-PI Scale	Time 1 Attriters		Time 2 Attriters			
	M	*SD*	*M*	*SD*	*t*	*df*
BOLDS Analog Neuroticism	98.23	14.72	95.26	16.51	-1.36	203
BOLDS Analog Extraversion	119.74	12.13	119.48	13.38	-0.14	196
BOLDS Analog Openness	112.25	12.35	110.66	13.76	-0.87	202
BOLDS Analog Agreeableness	105.33	8.47	106.14	8.63	0.69	207
BOLDS Analog Conscientious.	112.28	15.05	115.49	13.00	1.52	175

Note. "Time 1 Attriters" refers to cadets who left USMA prior to completing their first academic semester ($N = 77-101$), and "Time 2 Attriters" refers to those who left anytime after their first semester ($N = 100-108$). No *t*-values were statistically significant.

Conclusions

The initial objective of this analysis was to determine whether the analog scales developed by Evans would show similar relationships among the BOLDS data as they had among the data from which they were empirically developed. Such similarity would suggest that the results are not sample-specific and, hence, the analog scales are independently useful in other data sets.

Overall, the relationships between the analog scales and criterion measures emerging from the BOLDS data were similar to those described by Evans (1997). Regarding the correlations between analog ABLE/NEO-PI scale scores and cadets' mean leadership grades, 20 out of 24 comparisons showed no significant difference between the correlations resulting from the initial data and those from BOLDS. Among the four differences that emerged, three indicated that the correlation between the analog scale score (ABLE Dominance, ABLE Emotional Stability, ABLE Total) and the MD grade was stronger within the BOLDS data than it had been in Evans' original work.

Comparing the mean analog scale scores of USMA graduates and non-graduates resulted in three "discrepancies" with the previous analog results and nine "confirmations." Among the three differences, in two cases (ABLE Dominance and NEO-PI Extraversion) BOLDS data revealed no difference between the mean scale scores of graduates and non-graduates, whereas Evans found differences using both the original scales and the analogs. In the third case (NEO-PI Agreeableness), the mean scale scores of graduates and non-graduates differed in the BOLDS data, unlike in the Class of 1996 results.

The fact that the current results closely replicate those found by Evans (1997) suggests these results are generalizable – to other classes of USMA cadets, anyway – and these empirically derived analog scales are meaningful outside of their original context. Their meaning as temperament measures stems from their ability to reproduce the relationships that would have occurred had the actual ABLE or NEO-PI been administered.

Though the results of this investigation seem promising, there were limitations on these analyses that must be considered. First, the mean leadership grades in the BOLDS data were not computed in exactly the same manner as they had been with the earlier data sets. Specifically, the mean grades from BOLDS were calculated on seven (not all eight) academic semesters because the final senior year grades had not yet been included in ARI's data set. Similarly, the BOLDS analyses used mean grades from five (not four) summer details because there was no explanation in Evans (1997) to indicate which summer details had been included. A second limitation resulted from the classification of USMA graduates and non-graduates based on BOLDS data. According to the BOLDS database, 284 cadets did not receive military development grades through the seventh semester, suggesting they were no longer enrolled. However, according to USMA's final records, 883 cadets graduated in the Class of 1998, meaning that only 260 did not graduate.

Nonetheless, because the results of the current analyses so closely resembled those of Evans (1997), use of these ABLE and NEO-PI analog scales in future BOLDS analyses seems warranted.

PART II: REPLICATING MAEL AND WHITE'S (1994) ABLE-RELATED ANALYSES

       The Army's Assessment of Background and Life Experiences (ABLE) was administered to the USMA Class of 1994 as part of a project conducted by Mael and White (1994) to determine whether biodata scales could parallel temperament scales in their relationship to cadet leadership performance. Additionally, they examined whether either the overall ABLE score or BioABLE (the biodata equivalent) added incremental validity over and above the measure currently used by USMA to assess leadership potential. In light of their results, Mael and White reported, "It is possible to develop objective biodata measures that will be substantially analogous to valid temperament measures" (p. 306). Using BOLDS data to replicate their analyses, the current research compared the results achieved by using the *analog* ABLE scales with the results cited by Mael and White in association with the *actual* ABLE scales. The goal was to determine whether the analog ABLE scales perform similarly to the actual ABLE scales, thus suggesting their fidelity as temperament measures.

Method

Participants

USMA Class of 1994

       In the summer of 1990, 1,325 incoming cadets in the Class of 1994 completed questionnaires for Mael and White's biodata study. Of this group, 88% were male.

USMA Class of 1998

       In the summer of 1994, 1,143 incoming cadets in the Class of 1998 completed questionnaires. As part of the BOLDS research, their SIF and CCI survey items were used to calculate analog ABLE scores based on Evans' (1997) empirically derived scales. Of these cadets, 87% were male.

Measures

ABLE Scales

       Mael and White (1994) used an 88-item version of the ABLE, which included the following six scales: Dominance, Energy Level, Work Orientation, Emotional Stability, Dependability (composed primarily of items dealing with endorsement of traditional values), and Social Desirability.

       They also developed a 73-item biographical data questionnaire, with items intended to be objective, first-person, and, in principle, verifiable. By empirically keying these biodata items to the ABLE (for methodological details, see Mael, 1991, and Mael and Schwartz, 1991), they created biodata analogs of five ABLE scales: Bio-Dominance, Bio-Energy Level, Bio-Work

Orientation, Bio-Emotional Stability, and Bio-Dependability. In addition, they constructed a biodata composite for the whole ABLE, known as BioABLE.

As noted in Part I of this report, the regression equations Evans (1997) developed to compute analog scores for the ABLE scales appear in Appendix A. The variables included in the equations refer to SIF or CCI items.

External Criteria

Mael and White (1994) correlated ABLE and BioABLE scale scores with three criterion measures: Cadet Performance Reports (CPRs) completed by peers and superiors in relation to (1) Cadet Basic Training (CBT), which occurs during the summer preceding their freshman year (first summer); (2) the cadets' first academic term at USMA (Fall); and (3) Cadet Field Training (CFT), which occurs during the summer following their freshman year (second summer).

The CPR is a rating instrument consisting of 12 leadership dimensions: namely, duty motivation, military bearing, teamwork, influencing others, consideration for others, professional ethics, planning and organizing, delegating, supervising, developing subordinates, decision-making, and oral and written communication. A rated cadet is given a score of 1 ("needs much improvement") to 5 ("excellent") on each dimension. In addition, two dimensions are selected as the cadet's relative strengths and two as his/her relative weaknesses. Lastly, cadets are given an overall ranking that indicates whether their performance is in the upper 10%, upper 25%, middle 30%, lower 25%, or lower 10% of cadets in that particular duty position.

CPRs are primarily completed by cadets, from superior, peer, and subordinate positions. Cadet leaders are required to complete CPRs on those in subordinate duty positions. Peer and subordinate raters, on the other hand, nominate and rate the cadets having the highest and lowest performance relative to others in a particular group (for details, see Schwager & Evans, 1996). Academic instructors follow similar procedures, by nominating and rating the highest and lowest performers enrolled in their courses. Hence, it is possible for any cadet to receive multiple peer, subordinate, and instructor CPRs or none at all, depending on whether he/she is considered a high, low, or average performer.

In Mael and White's (1994) correlational analysis, they used a weighted combination of the ratings given by peers and superiors to the CPR's 12 dimensions of cadet performance. Because specific information about their "weighted combination" was not available, the computation of these ratings could not be replicated exactly for the current analyses. In place, *similar* correlations resulting from the BOLDS' analog ABLE scales were computed.

Specifically, in the BOLDS data, composite scores for the CPRs were computed by taking the mean of the overall ranking and two dimensions (i.e., duty motivation and military bearing) across all rater types (e.g., superior, peer, instructor). This formula originated from the work of Tisak (2000), who found that of the 12 dimensions, "duty motivation" and "military bearing" were most strongly related to the cadet's overall ranking, across rater types. He suggested this formula for a CPR composite score as a way of strengthening the overall ranking

measure while compensating for "missing" CPRs (i.e., not all cadets received CPRs from all types of raters).

Because the computation of CPR ratings in BOLDS differed from that in Mael and White (1994), an additional correlation was calculated for comparison purposes between the analog ABLE scales and the cadets' leadership (or "military development"; MD) grade for CBT, the fall term, and CFT. (For a description of military development grades, refer to the "External Criteria" section under "Measures" in Part I of this report.)

The correlations among the three criterion measures from 1994 and the six measures from 1998 are shown in Table 10. The CPRs from different sessions (i.e., CBT, Fall, CFT) and the MD grades from different sessions were similarly correlated (.23 to .41). Mael and White (1994) reasoned that because the three measures from 1994 were only moderately correlated with one another and they had "differential relationships with the predictors," they should not be combined or treated as repeated measures of the same criterion.

Table 10.
Correlations Among Criterion Measures for USMA Classes of 1994 and 1998

Criterion	1994 CBT CPR	1994 Fall CPR	1998 CBT CPR	1998 Fall CPR	1998 CFT CPR	1998 CBT MD	1998 Fall MD
'94 CBT CPR							
'94 Fall CPR	.35						
'94 CFT CPR	.39	.23					
'98 CBT CPR							
'98 Fall CPR			.33				
'98 CFT CPR			.33	.35			
'98 CBT MD			.76	.37	.35		
'98 Fall MD			.37	.56	.28	.38	
'98 CFT MD			.36	.32	.71	.41	.28

Note. All correlations of 1998 measures were significant at *p* < .001.

Results

Mael and White's (1994) correlational analysis is shown in Table 11, along with the respective correlations from BOLDS (for both the CPR and the MD grade). Similar to the results for the *actual* ABLE scales reported by Mael and White, in most cases, the BOLDS analog ABLE scales were related to leadership performance. The few non-significant relationships among the BOLDS data were as follows: (1) analog Dominance was not related to the CPR composite score from the cadets' first academic term, (2) analog Work Orientation and analog Emotional Stability were not related to the leadership (MD) grade from the first term, and (3) analog Work Orientation was not related to the MD grade from the second summer session.

Table 11.
Correlations Between ABLE Scales and Criterion Measures for
USMA Classes of 1994 and 1998

Variable	Criterion	First Summer (CBT)	First Term (Fall)	Second Summer (CFT)
'94 Dominance	CPR	.12***	.07**	.17***
'94 Bio-Dominance	CPR	.07*	.06*	.20***
'98 BOLDS Analog Dominance	CPR	.10**	.07	.09*
	MD	.12***	.11**	.12***
'94 Energy Level	CPR	.18***	.06*	.20***
'94 Bio-Energy Level	CPR	.16***	.08**	.23***
'98 BOLDS Analog Energy Level	CPR	.14***	.15***	.09*
	MD	.12***	.10**	.12**
'94 Work Orientation	CPR	.10***	.14***	.10***
'94 Bio-Work Orientation	CPR	-.01	.12***	.17*
'98 BOLDS Analog Work Orientation	CPR	.14***	.12***	.10**
	MD	.13***	.06	.06
'94 Emotional Stability	CPR	.17***	.03	.14***
'94 Bio-Emotional Stability	CPR	.17***	.06*	.23***
'98 BOLDS Analog Emotional Stab.	CPR	.09*	.10**	.09*
	MD	.12**	.07	.14***
'94 Dependability	CPR	.10***	.14***	.08*
'94 Bio-Dependability	CPR	-.01	.13***	.05
'98 BOLDS Analog Traditional Values	CPR	.15***	.16***	.14***
	MD	.11**	.11**	.14***
'94 ABLE Total	CPR	.19*	.11***	.20***
'94 BioABLE	CPR	.07*	.11***	.19***
'98 BOLDS Analog Total	CPR	.13***	.14***	.11**
	MD	.14***	.10*	.16***

Note. Class of 1994: *N* = 1183 (CBT and Fall); *N* = 1076 (CFT). Class of 1998 (BOLDS): *N* = 636-854 (Summer 1994); *N* = 634-818 (Fall); *N* = 612-778 (Summer 1995).
*p < .05. **p < .01. ***p < .001.

The BOLDS analog scales that most strongly related to the criterion measures differed from the actual ABLE scales Mael and White (1994) found to be most strongly related. They reported that Energy Level, Emotional Stability, and Dominance had the highest correlations with leadership performance during CBT and CFT, and Work Orientation and Dependability had the highest correlations with fall ratings. Among BOLDS data, in relation to CPR scores across all three sessions (i.e., first summer, fall, and second summer), analog Traditional Values, analog Work Orientation, and the analog ABLE Total had consistently strong relationships. With respect to the MD grades, analog Dominance, analog Energy Level, and analog Traditional Values had the strongest relationships consistently across all three sessions. In relation to MD grades from the summer sessions only (i.e., CBT and CFT), Analog ABLE Total and analog Emotional Stability had the strongest correlations.

Another goal of Mael and White's (1994) research was to determine whether the temperament and biodata scales contributed to the prediction of cadet leadership performance beyond that predicted by USMA's admissions measures. USMA uses the following entrance measures, among others, to predict cadet success:
- Scholastic Aptitude Test (SAT) score;
- High school rank;
- Physical Aptitude Exam (PAE) score;
- Leader Potential Score (LPS; which is computed by summing three ratings based on a cadet's high school record: 0.33 Extracurricular Activities Score, 0.33 Athletic Activities Score, and 0.33 Faculty Assessment Score); and
- Whole Candidate Score (WCS; computed by summing 0.6 College Entrance Examination Rating [based on SAT or ACT score and high school rank], 0.3 LPS, and 0.1 PAE).

Correlations between these five leadership predictors and BOLDS analog ABLE Total are listed in Table 12. As shown, analog ABLE Total was most highly correlated with LPS ($r = .14$, $p < .001$) and WCS ($r = .11$, $p < .01$). WCS was highly correlated with high school rank ($r = .64$, $p < .001$), SAT score ($r = .52$, $p < .001$), and LPS ($r = .46$, $p < .001$), not surprisingly, given it is directly computed from these scores. Interestingly, SAT score was negatively correlated with PAE ($r = -.20$, $p < .001$) and LPS ($r = -.19$, $p < .001$), suggesting an inverse relationship between cognitive aptitude and athletic abilities.

Table 12.
Correlations Among USMA Leadership Predictors and 1998 Analog ABLE Total

Variable	1. Analog ABLE Tot.	2. WCS	3. SAT	4. H.S. Rank	5. LPS
1. Analog ABLE Total					
2. WCS	.11**				
3. SAT	.00	.52***			
4. High school rank	.08*	.64***	.11***		
5. LPS	.14***	.46***	-.19***	.18***	
6. PAE	.09*	.08**	-.20***	-.09**	.11***

Note. $N = 690$-1142.
*$p < .05$. **$p < .01$. ***$p < .001$.

These correlations are comparable to those cited by Mael and White (1994), with the following three exceptions: (1) the correlation between WCS and PAE was negative in 1994 ($r = -.11$, $p < .001$) and positive in 1998 ($r = .08$, $p < .01$), (2) the correlation between SAT and high school rank was non-significant in 1994 (-.03) but significant and positive in 1998 ($r = .11$, $p < .001$), and (3) the correlations of WCS with SAT and LPS were stronger in 1998 ($r = .52$ and $r = .46$, $p < .001$, respectively) than in 1994 ($r = .21$ and $r = .12$, $p < .001$, respectively) (Fisher's z's = -8.7 and -9.3, $p < .001$, respectively). Nonetheless, the correlations associated with the 1998 *analog* ABLE Total (second column of Table 12) closely mimic those associated with the 1994 *actual* ABLE Total (see middle two columns of Table 13).

Table 13.
Correlations Between USMA Predictors and 1998 Analog ABLE Total,
1994 ABLE Total, and 1994 BioABLE

Variable	1998 Analog ABLE Total	1994 ABLE Total[1]	1994 BioABLE[1]
1. WCS	.11**	.11***	.24***
2. SAT	.00	.01	-.06*
3. High school rank	.08*	.09***	.22***
4. LPS	.14***	.13***	.37***
5. PAE	.09*	.12***	.07*

[1] Source: Mael and White, 1994.
*$p < .05$. **$p < .01$. ***$p < .001$.

Correlations between the predictors and the leadership criteria from the first summer, the first academic term, and the second summer are presented in Table 14 for the classes of 1994 and 1998. Across both classes, the LPS was related to the criteria from all three sessions (i.e., CBT, fall term, and CFT). Generally, the WCS, SAT score, and high school rank were related to the criteria from the first academic term but not from the summer sessions (i.e., CBT or CFT). Conversely, the PAE score was related to leadership criteria from the summer sessions (CBT and CFT) but not from the academic term. These findings reiterate the notion that leadership performance measured during the field exercises (CBT and CFT) differs from that measured during the academic terms.

Table 14.
Correlations Between USMA Predictors and Criterion Measures for
USMA Classes of 1994 and 1998

Predictor Variable	Criterion	First Summer (CBT)	First Term (Fall)	Second Summer (CFT)
'94 WCS	CPR	.04	.20***	.06
'98 WCS	CPR	.12***	.26***	.17***
	MD	.12***	.22***	.06
'94 SAT	CPR	.01	.02	-.06
'98 SAT	CPR	-.04	.11**	-.01
	MD	-.03	.11***	-.04
'94 High School Rank	CPR	-.01	.17***	-.03
'98 High School Rank	CPR	.05	.18***	.09**
	MD	.03	.09**	-.02
'94 LPS	CPR	.06*	.07*	.15***
'98 LPS	CPR	.11***	.09**	.17***
	MD	.10***	.12***	.11***
'94 PAE	CPR	.17***	.02	.19***
'98 PAE	CPR	.18***	.05	.12***
	MD	.23***	.06	.16***

*$p < .05$. **$p < .01$. ***$p < .001$.

Following Mael and White's work, hierarchical multiple regressions were performed to determine the incremental contribution of analog ABLE Total over and above the WCS (USMA's primary selection measure) in accounting for variance in the criterion measures. In other words, would the analog ABLE Total scale help USMA to predict the future success of its cadets?

For the CPR criterion measure from each of the three sessions (CBT, first term, and CFT) for the Class of '98, analog ABLE Total provided incremental validity over the WCS (see Table 15). In contrast, results varied by session when regressing the MD criterion measure. For the MD from CBT, analog ABLE Total added incremental validity over the contribution of WCS, and, in fact, with analog ABLE Total in the equation, the contribution of WCS became nonsignificant. On the other hand, analog ABLE Total did not account for significant variance in the MD from the first academic term, and WCS did not account for significant variance in the MD from CFT. Mael and White's results were similar, in that ABLE and BioABLE added incremental validity over WCS in relation to the criterion measure from each of the three sessions, and WCS did not account for significant variance in the criterion measure from CFT when entered with either the ABLE or BioABLE.

Table 15.
Hierarchical Regression Results Depicting the Incremental Validity of
Analog ABLE Total Over and Above the WCS

Session & Criterion	Step & Item	F	ΔR^2	Adjusted R^2	N
First Summer (CBT)					
CPR	1. WCS	9.41**	.013	.012	
	2. Analog ABLE Total	9.67***	.014	.024	693
MD	1. WCS	5.50*	.008	.006	
	2. Analog ABLE Total	9.15***	.017	.022	723
First Term (Fall)					
CPR	1. WCS	47.33***	.065	.063	
	2. Analog ABLE Total	28.23***	.012	.073	688
MD	1. WCS	35.62***	.049	.048	
	2. Analog ABLE Total	19.52***	.005	.051	695
Second Summer (CFT)					
CPR	1. WCS	16.44***	.024	.023	
	2. Analog ABLE Total	11.26***	.009	.030	665
MD	1. WCS	1.22	.002	.000	
	2. Analog ABLE Total	8.55***	.023	.022	667

Mael and White (1994) also examined the relationship of ABLE scales and USMA predictors to socially desirable responding (see Table 16 for a comparison of the correlations associated with the analog and actual ABLE Social Desirability scales). Moreover, they found the ABLE Social Desirability score was unrelated to the CBT and fall leadership ratings and only slightly related to the second summer ratings ($r = .06$, $p < .05$). In contrast, the BOLDS analog ABLE Social Desirability score was related to both the CPR composite and the MD grade from the first summer; the CPR composite from the fall term; and, though less strongly, the CPR composite and MD grade from the second summer (as shown in Table 16).

Table 16.
Correlations of ABLE Scales, USMA Predictors, and Criterion Measures With
Social Desirability (Validity) Scales

ABLE Scale/ USMA Predictor/ Criterion Measure	1998 Analog ABLE Social Desirability	1994 ABLE Social Desirability[1]
Dominance	.21***	.09***
Energy Level	.48***	.23***
Work Orientation	.65***	.39***
Emotional Stability	.42***	.16***
Dependability/ Traditional Values	.74***	.31***
ABLE Total	.56***	.34***
WCS	.09*	.04
SAT	-.06	-.03
High school rank	.15***	.09***
LPS	.10**	.06***
PAE	.05	-.00
CPR: CBT	.17***	ns[2]
Fall	.14***	ns[2]
CFT	.08*	.06*
MD: CBT	.14***	–
Fall	.07	–
CFT	.08*	–

Note. Class of 1998 (BOLDS): $N = 677\text{-}801$. Class of 1994: $N = 1314\text{-}1334$.
Correlations with the MD grade are not applicable to the 1994 data.
[1] Source: Mael and White, 1994.
[2] ns = nonsignificant correlation, magnitude unknown.
*$p < .05$. **$p < .01$. ***$p < .001$.

Conclusions

The results of this replication exercise are encouraging. When the individual analog ABLE scale scores were correlated with CPR ratings and MD grades, each of the scales was shown to relate to cadet leadership performance, similar to Mael and White's finding. Moreover, when the 1998 analog ABLE Total score was correlated with USMA predictors of leadership potential (e.g., WCS, SAT, LPS), the resulting correlations closely resembled those attained from the 1994 actual ABLE Total. Lastly, hierarchical regression results indicated that the analog ABLE Total score tended to provide incremental validity over and above the Whole Candidate Score that USMA has used to predict leadership performance and cadet success, just as Mael and White had found.

One goal of Mael and White's work was to determine if ABLE motivational constructs could be measured with less personally intrusive and less fakable biodata items. They analyzed the relative relationship of the ABLE and biodata scales, USMA predictors, and criterion measures to socially desirable responding. Using 1998 data to replicate these correlations resulted in one major difference. Whereas Mael and White reported that socially desirable responding did not account for significant variance in two of the three criteria, 1998 data showed the analog ABLE Social Desirability scale tended to be positively correlated with the criteria. Mael and White explained that in previous studies where socially desirable responding (or faking) was positively related to criteria, it was interpreted as a measure of self-esteem and treated as meaningful variance rather than measurement error (Zerbe & Paulhus, 1987). Whether this interpretation holds for an empirically derived analog Social Desirability scale within the BOLDS data is unclear. Nonetheless, because we do not know the extent to which the computation of criterion measures differed in 1998 compared to 1994, disparities in resulting relationships may not be surprising.

Aside from the criterion-related nature of the analog Social Desirability scale, other results associated with the BOLDS analog ABLE scales closely replicated the findings Mael and White (1994) reported in conjunction with the actual ABLE scales. The similarity in findings adds credence to Evans' analog scales as valid substitutes for temperament measures.

DISCUSSION

The analyses described in this report demonstrate that the ABLE and NEO-PI analog scales performed similarly with the BOLDS data as they had with the data from which they were developed. This consistency suggests the results were not sample-specific and, hence, the analog scales produced reliable scores among cadets. Moreover, the analog ABLE scales performed comparably to the *actual* ABLE scales from the Class of '94, indicating a degree of credibility – if not convergent validity – for these empirically derived scales.

One of the greatest limitations of such replication analyses hinges on how accurately the measures can be duplicated. In reproducing both Evans' (1997) and Mael and White's (1994) analyses, the computation of leadership performance criterion measures could not be exactly

replicated, due to lack of information. Even so, the similarity of the results suggests the robustness of the relationships that emerged.

The overall goal of examining these analog scales was to provide insight into their validity as temperament measures and their usefulness in future BOLDS analyses. Being scales empirically developed from one data set, their legitimacy outside of their original context was uncertain. Because similar results emerged from rerunning analyses using new data, this investigation demonstrates, in a sense, the soundness of these analog measures. With respect to BOLDS, the soundness of the analog scales is important because they are the only personality measures collected on cadets in the Class of 1998 while they were at West Point. As temperament has been shown to be linked to certain leadership behaviors and effectiveness (Hogan, Curphy, & Hogan, 1994), having such measures in BOLDS is valuable. Based on the results of this investigation, it is recommended that in future BOLDS analyses regarding leadership potential and performance, consideration should be given to the possible contribution the analog ABLE and NEO-PI scales could make.

According to analyses reported in Part I, none of the ABLE or NEO-PI scales *alone* would be effective predictors of cadet leadership performance, though Conscientiousness, Work Orientation, and Social Desirability showed consistent weak positive correlations with cadets' leadership grades. Nor would any of the scales *alone* make successful predictors of cadet attrition at USMA, given the relatively small magnitude of differences between the scale scores of USMA graduates and those of non-graduates. Nonetheless, these findings do not preclude the potential indirect effects of single temperament scales, the impact of clusters of traits, or the predictive value of these scales on future assessments of leadership performance.

Aside from BOLDS, the analog ABLE and NEO-PI scales might be valuable to other USMA research. Considering that the scales are based on survey items administered annually to incoming cadets, the measures are relatively cost effective – that is, until many of the items become obsolete. Even then, Evans' methodology could be adapted to new survey items. This measurement procedure saves both cadet time and the cost of personality instruments, while potentially contributing a significant dimension to archival data sets. Moreover, as empirically derived analogs, the scales may be more resistant to socially desirable responding and faking. As a first step toward verifying the usefulness of Evans' (1997) analog ABLE and NEO-PI scales, this investigation produced promising results. It is incumbent on future research, however, to focus on establishing the construct validity of these analog scales.

REFERENCES

Astin, A. W., Korn, W. S., & Berz, E. R. (1990). *The American freshman: National norms for Fall 1990*. Los Angeles: Higher Education Research Institute, UCLA.

Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.

Evans, K. L. (1997, April). *Estimating personality constructs from archival data* (ARI Technical Report 1063). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A328 816)

Hogan, R., Curphy, G. J., & Hogan, J. (1994). What we know about leadership: Effectiveness and personality. *American Psychologist, 49*(6), 493-504.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities [Monograph]. *Journal of Applied Psychology, 75*(5), 581-595.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspective. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford Press.

Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology, 44*, 763-792.

Mael, F. A., & Hirsch, A. C. (1993). Rainforest empiricism and quasi-rationality: Two approaches to objective biodata. *Personnel Psychology, 46*, 719-738.

Mael, F. A., & Schwartz, A. C. (1991). *Capturing temperament constructs with objective biodata* (ARI Technical Report 939). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A245 119)

Mael, F. A., & White, L. A. (1994). Motivated to lead: Dispositional and biographical antecedents of leadership performance. In H. F. O'Neil, Jr., & M. Drillings (Eds.), *Motivation: Theory and Research*, (pp. 285-311). Hillsdale, NJ: Lawrence Erlbaum Associates.

Milan, L. M., Bourne, Jr., D. R., Zazanis, M. M., & Bartone, P. T. (in press). *Measures collected on the USMA class of 1998 as part of the Baseline Officer Longitudinal Data Set (BOLDS)* (ARI Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Schwager, E. H., & Evans, K. L. (1996, April). *An Exploration of the Construct Validity of a Leadership Behavior Rating System* (ARI Technical Report 1041). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A310 259)

White, L. A., Nord, R. D., Mael, F. A., & Young, M. C. (1993). The Assessment of Background and Life Experiences (ABLE). In T. Trent & J. H. Laurence (Eds.), *Adaptability screening for the Armed Forces* (pp. 101-162). Washington, DC: Office of Assistant Secretary of Defense (Force Management and Personnel).

Zaccaro, S. J. (1996, October). *Models and theories of executive leadership: A conceptual/empirical review and integration*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review, 12*(2), 250-264.

APPENDIX A: PREDICTION EQUATIONS FOR SEVEN ABLE SCALES

Analog Dominance = 0.90 + 0.23(self-rated leadership ability) + 0.06(self-rated social self-confidence) + 0.04(importance of having administrative responsibility for the work of others) + 0.05(self-rated public speaking ability) + 0.02(importance of influencing the political structure) + 0.04(self-rated competitiveness) - 0.02(self-rated emotional health) + 0.03(importance of becoming an authority in one's field) + 0.02(chances of election to a student office) - 0.02(self-rated intellectual self-confidence)

Analog Energy Level = 1.39 + 0.08(self-rated drive to achieve) - 0.06(felt depressed during the past year) + 0.07(chances of being satisfied with one's college) + 0.08(self-rated physical health) - 0.04(importance of leadership training, as a reason for seeking appointment to USMA) - 0.07(felt overwhelmed during the past year) + 0.03(self-rated leadership ability) - 0.03(chances of dropping out permanently) + 0.03(self-rated public speaking ability) + 0.02(self-rated competitiveness) + 0.02(self-rated emotional health)

Analog Work Orientation = 1.32 + 0.18(self-rated drive to achieve) + 0.05(hours spent in a typical week studying or doing homework during last year in high school) - 0.13(failed to complete a homework assignment on time during the past year) + 0.06(importance of being made more cultured, in deciding to go to college) - 0.06(self-reported academic rank in high school graduating class) - 0.08(was bored in class during the past year) + 0.04(importance of becoming an authority in one's field) + 0.04(self-rated physical health) + 0.03(self-rated cooperativeness)

Analog Emotional Stability = 2.08 - 0.13(felt depressed during the past year) + 0.07(self-rated emotional health) - 0.12(felt overwhelmed during the past year) + 0.06(chances of being satisfied with one's college) + 0.04(self-reported social self-confidence) - 0.04(chances of dropping out permanently) - 0.02(present career intention) + 0.02(self-rated physical health) + 0.02(self-rated understanding of others)

Analog Traditional Values = 2.47 + 0.06(chances of being satisfied with one's college) - 0.08(view that marijuana should be legalized) - 0.04(desire to be an Army officer, as a reason for seeking appointment to USMA) - 0.06(drank beer during the past year) + 0.09(choice of present college) - 0.05(failed to complete a homework assignment on time during the past year) + 0.04(self-rated cooperativeness) - 0.06(was bored in class during the past year) - 0.03(helpfulness of USMA catalog in college decision-making process) - 0.03(view that if two people really like each other, it's all right for them to have sex even if they've known each other for only a very short time) - 0.03(chances of dropping out permanently) + 0.01(hours spent in a typical week studying or doing homework during last year in high school) - 0.02(present career intention)

Analog Total = 1.71 + 0.06(self-rated leadership ability) + 0.05(chances of being satisfied with one's college) + 0.07(self-rated drive to achieve) - 0.05(felt depressed during the past year) - 0.07(was bored in class during the past year) - 0.03(chances of dropping out permanently) + 0.03(self-rated physical health) - 0.02(present career intention) + 0.02(self-rated emotional health) - 0.04(felt overwhelmed during the past year) + 0.02(self-rated public speaking ability) + 0.02(self-rated understanding of others) + 0.01(self-rated social self-confidence)

Analog Social Desirability = 1.77 - 0.07(failed to complete a homework assignment on time during the past year) + 0.06(self-rated cooperativeness) - 0.02(desire to be an Army officer, as a reason for seeking appointment to USMA) - 0.02(hours spent in a typical week socializing with friends during last year in high school) - 0.05(was bored in class during the past year) - 0.03(drank wine or liquor during the past year) - 0.04(came late to class during the past year) - 0.03(importance of being able to make more money, in deciding to go to college) - 0.02(chances of dropping out permanently) + 0.01(hours spent in a typical week studying or doing homework during last year in high school) - 0.03(felt depressed during the past year)

Note. The above weights represent the mean of two unstandardized parameter estimates, developed from random halves of a 727-cadet sample. Source: Evans, 1997.

Analog Neuroticism = 99.35 - 8.36(self-rated emotional health) + 8.11(felt depressed during the past year) + 1.57(likelihood of graduation from USMA) + 4.96(was bored in class during the past year) - 2.90(self-rated intellectual self-confidence) + 3.39(importance, in deciding to attend this college, of relatives wanting me to come here) - 2.74(chances of election to a student office) + 1.90(chances of changing career choice) + 2.86(felt overwhelmed during the past year) + 1.80(chances of failing one or more courses) - 1.05(self-rated public speaking ability)

Analog Extraversion = 33.36 + 5.43(self-rated popularity) + 3.82(self-rated social self-confidence) + 2.06(self-rated leadership ability) + 1.80(importance of becoming a community leader) + 1.71(importance of raising a family) + 1.91(chances of election to a student office) + 1.64(self-rated understanding of others) + 0.93(hours spent in a typical week partying during last year in high school) + 1.86(self-rated drive to achieve) + 1.18(self-rated public speaking ability) + 1.37(self-rated physical health)

Analog Openness to Experience = 52.21 + 5.43(importance of writing original works) + 2.28(importance of developing a meaningful philosophy of life) + 3.88(view that the Federal government is not doing enough to control environmental pollution) + 2.57(self-rated artistic ability) + 3.61(view that marijuana should be legalized) - 1.47(view that activities of married women are best confined to the home and family) + 2.97(importance of learning more about things that interest oneself, in deciding to go to college) + 3.29(importance of becoming accomplished in one of the performing arts) + 1.45(view that abortion should be legal) + 1.44(importance of helping to promote racial understanding) + 2.48(importance of gaining a general education and appreciation of ideas, in deciding to go to college) - 1.23(view that it is important to have laws prohibiting homosexual relationships) + 1.40(importance of helping others who are in difficulty)

Analog Agreeableness = 99.25 + 4.36(self-rated cooperativeness) - 3.05(importance of being very well off financially) + 3.24(importance of helping others who are in difficulty) - 2.53(drank wine or liquor during the past year) + 2.38(view that the death penalty should be abolished) - 2.23(view that an individual can do little to bring about changes in our society) + 1.91(view that colleges should prohibit racist or sexist speech on campus) - 2.40(view that there is too much concern in the courts for the rights of criminals) + 3.34(attended a religious service during the past year) - 1.07(hours spent in a typical week partying during last year in high school) - 2.72(was bored in class during the past year)

Analog Conscientiousness = 56.91 + 8.53(self-rated drive to achieve) - 2.33(present career intention) -5.11(failed to complete a homework assignment on time during the past year) + 2.77(self-rated emotional health) + 3.19(importance of having administrative responsibility for the work of others) - 4.46(came late to class during the past year) + 2.47(chances of being elected to an academic honor society) + 3.25(chances of being satisfied with one's college) - 2.69(felt depressed during the past year) + 1.52(self-rated cooperativeness) + 1.08(self-rated public speaking ability)

Note. The above weights represent the mean of two unstandardized parameter estimates, developed from random halves of a 635-cadet sample. Source: Evans, 1997.